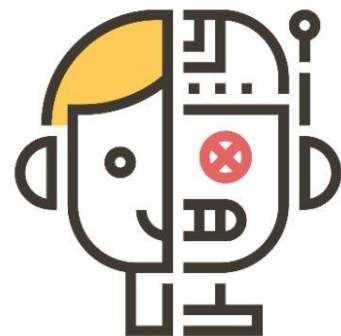


はじめての生成AI 【Level200】

チャットボットを理解するために 知っておきたいこと

2025年8月26日(火) 13:00



ブレインズテクノロジー株式会社



本セミナーの位置付け

生成AIの利用をお考えの方に向けた
3ステップのシリーズセミナーです



START

Level 100へ

Level 100

はじめての生成AI
~生成AIの利用を
検討するために
知っておきたいこと~

Level 200

はじめての生成AI
~チャットボットを
理解するために
知っておきたいこと~

Level 300

はじめての生成AI
~Chat EIの導入
評価検証について~

本セミナーの目的

チャットボットの製品を検討するために必要な基礎知識を習得する。
Chat EIについて紹介します。

【想定している参加者】

- Level 100に参加した方
- RAGの技術や検証例にご興味のある方
- 製品機能にご興味のある方



チャットボットの利用について社内で議論できるようになる



本日の内容

LLMの周辺技術

🔄めくる

Chat EIの
システム構成

🔄めくる

よくある課題と
検証事例

🔄めくる

セッション終了後にmiroボード上でご質問をお受けします

Chat EI - 会話型AIチャットシステムとは？

企業内の情報を学習基盤とした会話型AIチャットシステム
生成AI、生成AIの周辺技術、および、Neuron ESのクローリング技術を利用

クローリング

NeuronESで開発されたクローラーを利用可能

Embedding

単語や文章等をベクトル表現に変換すること

Transformer

それ以前に主流だった深層学習のアルゴリズムよりも高速かつ高精度で、自然言語処理やその他AI分野に革命的な進歩をもたらした生成AIのキー技術

RAG

(検索拡張生成)

検索ベースと生成ベースの両方の長所を組み合わせた自然言語処理技術

セマンティック検索

ベクトル検索のスコアに基づいて再順位付けするランカーを利用し、ハイブリッド検索の弱点を補完して最適な結果を提供

ケーブルが原因で発生した事故を教えてください

1. 事故報告書_19.docx
- 事故の状況: 当該事業場の高圧送込みケーブルが接触し、ケーブルが断線し、ケーブルが断線したため、屋内施設に接触して発生した。配電装置内のケーブルが断線し、波及事故に繋がったものと想定される。
- [出典] 事故報告書_19.docx
- [注釈部分の抜粋] 「当該事業場の高圧送込みケーブルが断線し、ケーブルが断線したため、屋内施設に接触した。屋内施設に接触して発生した。配電装置内のケーブルが断線し、波及事故に繋がったものと想定される。」

2. 事故報告書_16.docx
- 事故の状況: 高圧送込みケーブル(CV1997年製)の端末部分が断線が発生した。出火方式で事故が保護範囲内であったため、波及事故に繋がった。
- [出典] 事故報告書_16.docx
- [注釈部分の抜粋] 「高圧送込みケーブル(CV1997年製)の端末部分が断線が発生した。出火方式で事故が保護範囲内であったため、波及事故に繋がった。」

3. 事故報告書_17.docx
- 事故の状況: 高圧送込みケーブル(CV2011年製)が水トリー現象により断線不具合となり、断線が発生した。事故点は屋内施設保護範囲(DGR)の保護範囲内であったが、DGRが動作しなかったため、波及事故に繋がった。
- [出典] 事故報告書_17.docx
- [注釈部分の抜粋] 「高圧送込みケーブル(CV2011年製)が水トリー現象により断線不具合となり、断線が発生した。事故点は屋内施設保護範囲(DGR)の保護範囲内であったが、DGRが動作しなかったため、波及事故に繋がった。」

事故報告書_19.docx Preview

Message

ケーブルが原因で発生した事故を教えてください

過去の事故報告書では...

トラブル時の対応方法を教えてください

対応方法は次の3つがあります...

インターネット上の情報だけでなく、企業データを生成AIの学習基盤とする技術も登場

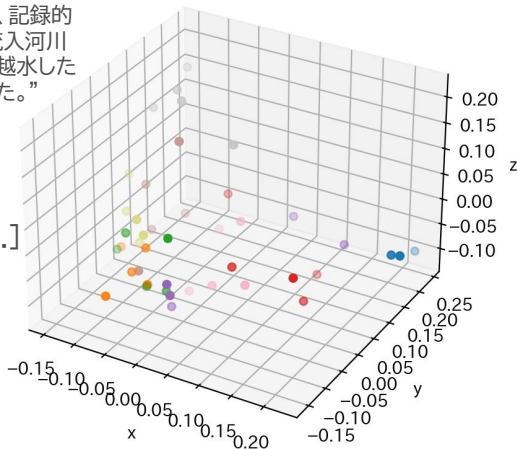
文章等

“当該水力発電所において、記録的な集中豪雨の発生により流入河川水が増加し、洪水吐き門を越水したため、異常放流事故になった。”

ベクトルに変換

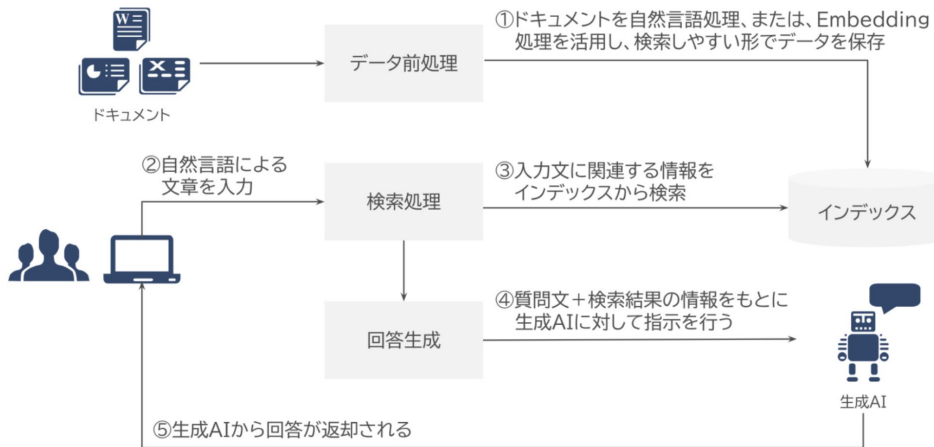
$[-0.131, 0.232, \dots]$

1536次元～
3072次元



Embedding

単語や文章等を
ベクトル表現に変換すること



RAG

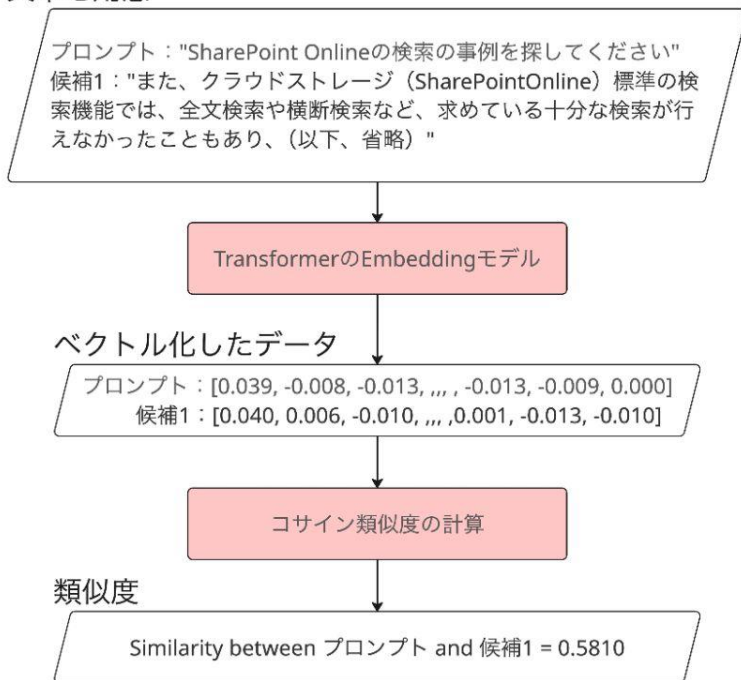
(Retrieval Augmented Generation)

検索ベースと生成ベースの両方の長所を組み合わせた自然言語処理技術

Embeddingとは

- LLMは内部でEmbedding技術を使用しています。具体的には、入力された単語や文書をまずEmbedding層でベクトル化し、その後トランスフォーマー構造で処理します。
- Embeddingは単独でも利用可能であり、大量のデータから関連性の高い情報を検索するための基盤技術としても活用されます。

文章を用意



python

```
1 from openai import AzureOpenAI
2
3 client = AzureOpenAI(azure_endpoint=endpoint,
4 api_version=version,
5 api_key=api_key)
6
7 def get_embedding(text):
8     response = client.embeddings.create(input=[text]
9     , model = model_id)
10    return response.data[0].embedding
11
12 embeddings = [get_embedding(text) for text in texts]
```

- Embeddingにより1536次元~3072次元のベクトルになります。
- EmbeddingのTransformerモデルは文脈に応じてEmbeddingを動的に変更できるため、単語やフレーズの意味を柔軟に捉えられます。
- 固定的なStatic Embeddingよりも精度が向上します。

実験

RAGの処理とエンタープライズサーチの処理の違い

comparison

	Embeddingを使うRAGの場合	エンタープライズサーチ
前処理 <small>実験</small>	<ul style="list-style-type: none">・ チャンク（最大チャンク長、オーバーラップ）・ 情報の削除（メールアドレス、URLなど）・ 文書の種類に応じた前処理（表形式の処理、QA形式に変換、など）・ Embedding（ベクトル化）	以下の処理を行っています。 <ul style="list-style-type: none">・ 形態素解析・ 2-Gram・ 各種フィルタ・ 各種変換
インデックス <small>(本で言うところの「目次」)</small>	～100ファイル	200万ファイル～1億ファイル以上
検索	類似度が高い順にデータをランキングし、上位k個の結果を返します。 top-kのパラメーターがあります。	キーワード検索で、ヒットしたファイル全てを結果表示します。選択したソート順に従って並びます。
回答	質問文と検索結果の上位k個の情報をもとにLLMに対して指示を行い、回答を得る。	生成AI連携の機能があれば、キーワード検索の上位k個の結果と質問文をもとにLLMに対して指示を行い、回答を得る。

チャンキングとは

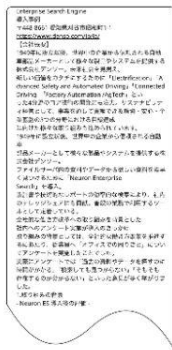
Word, Excel, PowerPoint, PDFから
プログラムでテキストを抽出するとい
くことをしています。
そのため、製品やサービスによって、
対応可能な拡張子が決まっています。

どのくらいの長さで小分けに
するか?のパラメーターが
あります。
(最大チャンク長、
オーバーラップ)

文書
(PDFファイル)



テキスト抽出



チャンク毎に分割



ベクトル化

[0.23, -0.12, -0.11, ...]

[-0.95, 0.62, -0.32, ...]

[0.95, 0.62, -0.32, ...]

ベクトルDBへ保存



セマンティック検索

検索クエリの意味や意図を解析し、より適切な情報を提供する技術です。キーワード検索よりも高度な技術で、質問文と回答に同一のキーワードが含まれていなくても、的確な結果を得られることが期待されます。

- ・ 従来型：ルールや辞書に基づく構文・意味解析が中心
- ・ AI型：ベクトル化・機械学習により文脈や意味全体を柔軟に捉える

特に、Transformerのリランクモデルを利用するセマンティック検索が提案されています。

★ ユーザーが検索窓に入力する語句や条件、質問文

ハイブリッド検索：ベクトル検索+キーワード検索

ベクトル検索

意味的な関連性に基づいて検索結果を提供します。曖昧なニュアンスや文脈を理解し、ユーザーのクエリに最も関連する情報を取得可能です。キーワードが一致しない場合でも、概念的に近い情報を検索できる柔軟性があります。

キーワード検索

正確な単語一致が必要な場合に有効です。例えば、製品コードや専門用語など、完全一致が重要な場面で優れた性能を発揮します。

リランクの利用：ハイブリッド検索+リランク

ベクトル検索

意味的関連性に基づいて広範囲の候補を取得しますが、曖昧なクエリや短いキーワードには弱い場合があります。

キーワード検索

単語一致に基づいて正確な結果を提供しますが、文脈や意味を十分に考慮できないことがあります。

リランク

関連性スコアに基づいて再順位付けすることで、両者の弱点を補完し、最適な結果を提供します。

実験

Level 100

RAGの中では
社内文書を
「検索」
する

🔄めくる

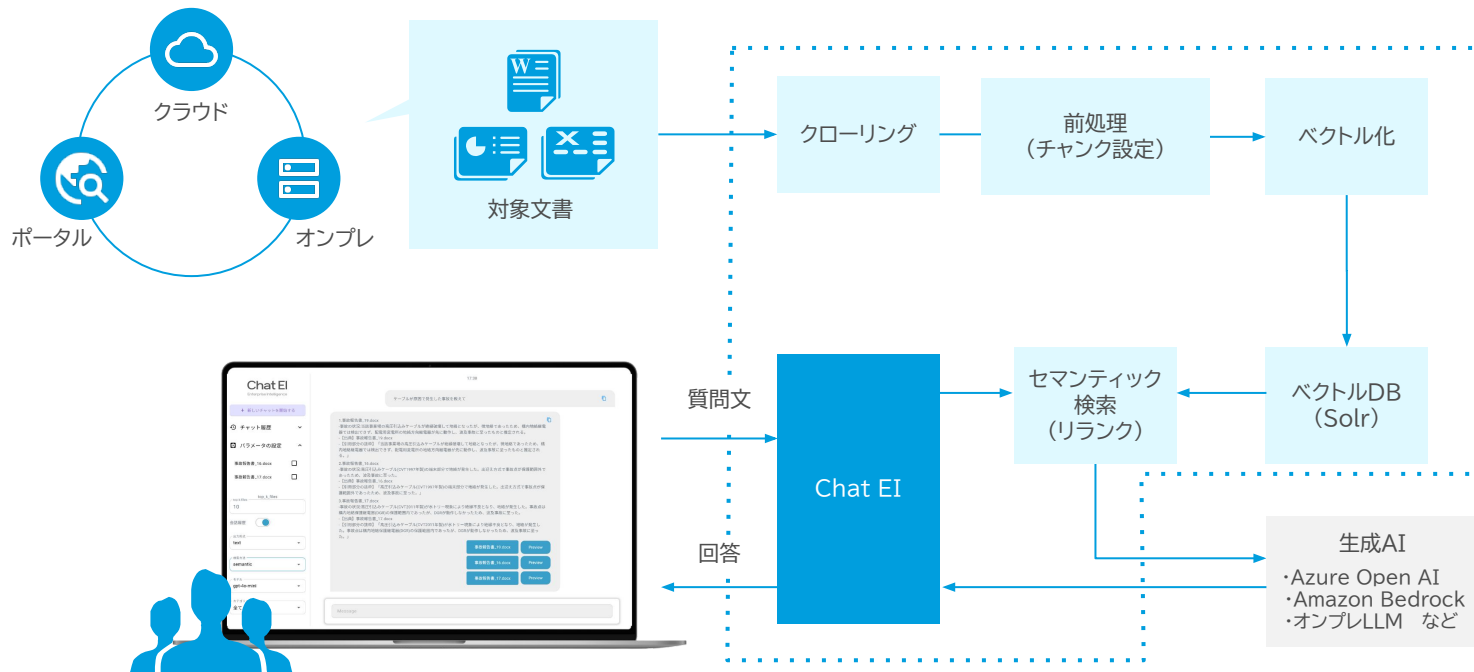
セマンティック検
索にはベクトル化
が必要です

🔄めくる

機械学習の基本は
「前処理」

🔄めくる

既存のNeuron Enterprise Searchの資産(クローリング技術)を活用し
生成AI系の機能を追加した形での提供が可能



会話履歴利用

入力した内容のみだけでなく
チャットでのやり取りも用いて
回答を生成します

ファイル指定

問合せ内容に近い検索結果の
ファイルを指定できます

ファイル数絞り込み

生成AIの回答で利用する対象
ファイル数の調整が行えます

出力フォーマット

テキスト・JSON・テーブルから
選択できます。表形式で回答結
果が欲しい場合はテーブルを選
択してください

AIモデルの指定

利用したいAIモデル(LLM)を
指定することが可能です

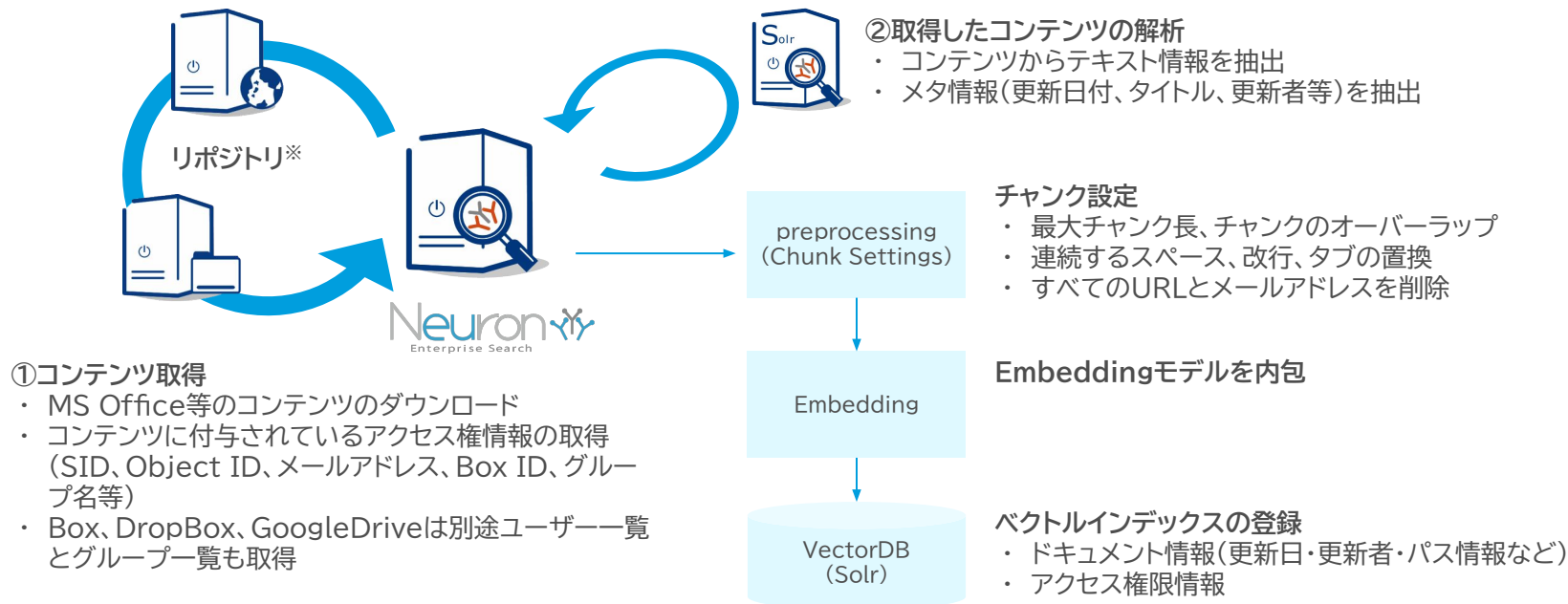
カテゴリー

Neuron ESで設定する取得元
での絞り込みを行うことが可能
です

The screenshot displays the Neuron ES search interface. On the left, a sidebar contains settings for chat history, parameter settings, file selection (top-k-files: 10), chat history toggle, output format (text), search method (semantic), AI model (gpt-4o-mini), and category (全て). The main area shows a search query: "ケーブルが原因で発生した事故を教えてください". Below the query, three search results are listed, each with a document title and a preview snippet. The results are: 1. 事故報告書_19.docx, 2. 事故報告書_16.docx, and 3. 事故報告書_17.docx. Each result has a "Preview" button. A callout box highlights the search method options: "hybridとsemanticから選択できます。 hybrid:キーワード検索とベクトル検索 semantic:ベクトル検索のみ". Another callout box highlights the document preview feature: "出典元・ファイル参照・プレビュー 出典元が回答内容に表示されます。また、出典ドキュメントへのプレビュー・参照を行うことができます".

機能	説明
認証・認可	導入環境に応じた認証機構と連携し、ユーザー認証を行うことができます。また、ドキュメントのアクセス権限に応じた検索結果の制限(認可)を行うことができます。これにより、アクセスできるドキュメントのみから回答を生成することが可能となります。
クローリング設定	Neuron ESで利用可能なレポジトリ・ファイル形式は、すべて利用することができます。 ※ 対応リポジトリ・解析対象ドキュメントを参照ください。
チャンク設定	文書から抽出したテキスト情報に対して、テキスト加工処理、最大チャンク長・オーバーラップ率の設定を行うことが可能です。
Embeddingモデルの変更	文章をベクトル化するモデルを変更して、インデックスを作成することが可能です。変更する際には、サーバー設定を変更してプロセスの再起動を行い、再度フルクローリングが必要となります。
セマンティック検索	セマンティック／ハイブリッド検索の検索方法の選択が可能です。今後は、リランキング手法への対応も視野に入れています。
生成AIモデルの変更	Azure OpenAI上にデプロイされた会話形式の生成AIモデルを設定することで、生成AIモデルの切り替えを行うことができます。(現状、共通設定となっておりプロセスの再起動が必要となります)

クローリング(リポジトリからコンテンツをダウンロードし
解析・インデックス作成する処理)をする際、下記をおこなっています



※ リポジトリ=ファイル保管場所のこと
(ファイルサーバー、Box、SharePoint Online、Dropbox、Google Drive等)

多くのお客様のシステム環境に対応すべく、検索対象を広げています

レポジトリ	接続先
<p>Notes</p>	<ul style="list-style-type: none"> Notes/Domino 6.5以降
<p>SharePoint SharePoint Online</p>	<ul style="list-style-type: none"> SharePoint 2013, 2016, 2019 https://xxxx.sharepoint.com
<p>box</p>	<ul style="list-style-type: none"> https://Box.com
<p>Dropbox</p>	<ul style="list-style-type: none"> https://www.dropbox.com/xxxx/
<p>Google Drive</p>	<ul style="list-style-type: none"> https://drive.google.com/xxxx/

レポジトリ	接続先
ファイルサーバ	<ul style="list-style-type: none"> SMB1, SMB2, SMB3(SMB暗号化)対応 ファイルサーバ(Windowsファイルサーバ, Samba, NetApp, NAS等)
ファイルシステム	<ul style="list-style-type: none"> Windowsファイルシステム(NTFS) Linux
Webサーバ	<ul style="list-style-type: none"> HTTP/HTTPS
データベース*	<ul style="list-style-type: none"> Oracle SQL Server

* データベースは一例です。他のデータベースをご希望の場合はお問い合わせください。
上記に掲載されていないレポジトリをご希望の場合はお問い合わせください。

分類	ファイル形式
テキストファイル	<ul style="list-style-type: none"> txt, text, csv等
MS Office文書	<ul style="list-style-type: none"> 【XP/2003形式】doc, xls, ppt 【2007以降 OOXML形式】docx, xlsx, pptx, ppsx, docm, xlsx, pptm, ppsm
Visio	<ul style="list-style-type: none"> 【2003/2007/2010形式】vsd*1 【2013以降 OOXML形式】vsdx
PDF	<ul style="list-style-type: none"> PDF文書
HTML	<ul style="list-style-type: none"> htm, html, xhtml, xhtm等 (HTML Version 1.0, 1.1, 2.0, Atom, RSS 2.0等)
リッチドキュメント	<ul style="list-style-type: none"> rtf形式のリッチテキスト
Open Document	<ul style="list-style-type: none"> odt, ods, odp, odg, ott, ots, otp, otg, sxw, sxc, sxi, sxd, stw, stc, sti, std形式のOpen Document

分類	ファイル形式
DocuWorks	<ul style="list-style-type: none"> xdw, xbd形式のDocuWorksドキュメント*2
圧縮ファイル	<ul style="list-style-type: none"> zip, bz2, tar, z, gz, tgz, 7zで圧縮された文書*3
マルチメディア (オーディオ、ビデオ、画像)	<ul style="list-style-type: none"> mp3, jpeg, mpeg等*4
メールファイル	<ul style="list-style-type: none"> msgファイル
CADファイル	<ul style="list-style-type: none"> dwg形式のAutoCADファイル
一太郎ファイル	<ul style="list-style-type: none"> 一太郎 ver5以降

- *1 2002以前のバージョンで作成されたファイルやテンプレートを利用しているファイルは解析できない場合があります
- *2 別途Fujixerox社のDocuWorks Viewer lightをインストールする必要があります
- *3 圧縮されたファイル内の文書のタイトルがインデクシングされます。標準では圧縮文書のコンテンツはインデクシングされません
- *4 ファイル名やサイズ、EXIF等のメタデータ(デフォルトでは検索対象外)のみ検索対象となります

検索システムで培った企業データのインデキシング技術、認証認可技術を使用した生成AIチャット

簡単

既存のデータストレージの
情報を学習基盤に

オンプレのファイルサーバや社内ポータル、クラウドストレージなどの資料やデータを学習基盤にできるので、学習用データのアップロード等の作業が不要

高精度

Embeddingによる
回答精度の向上

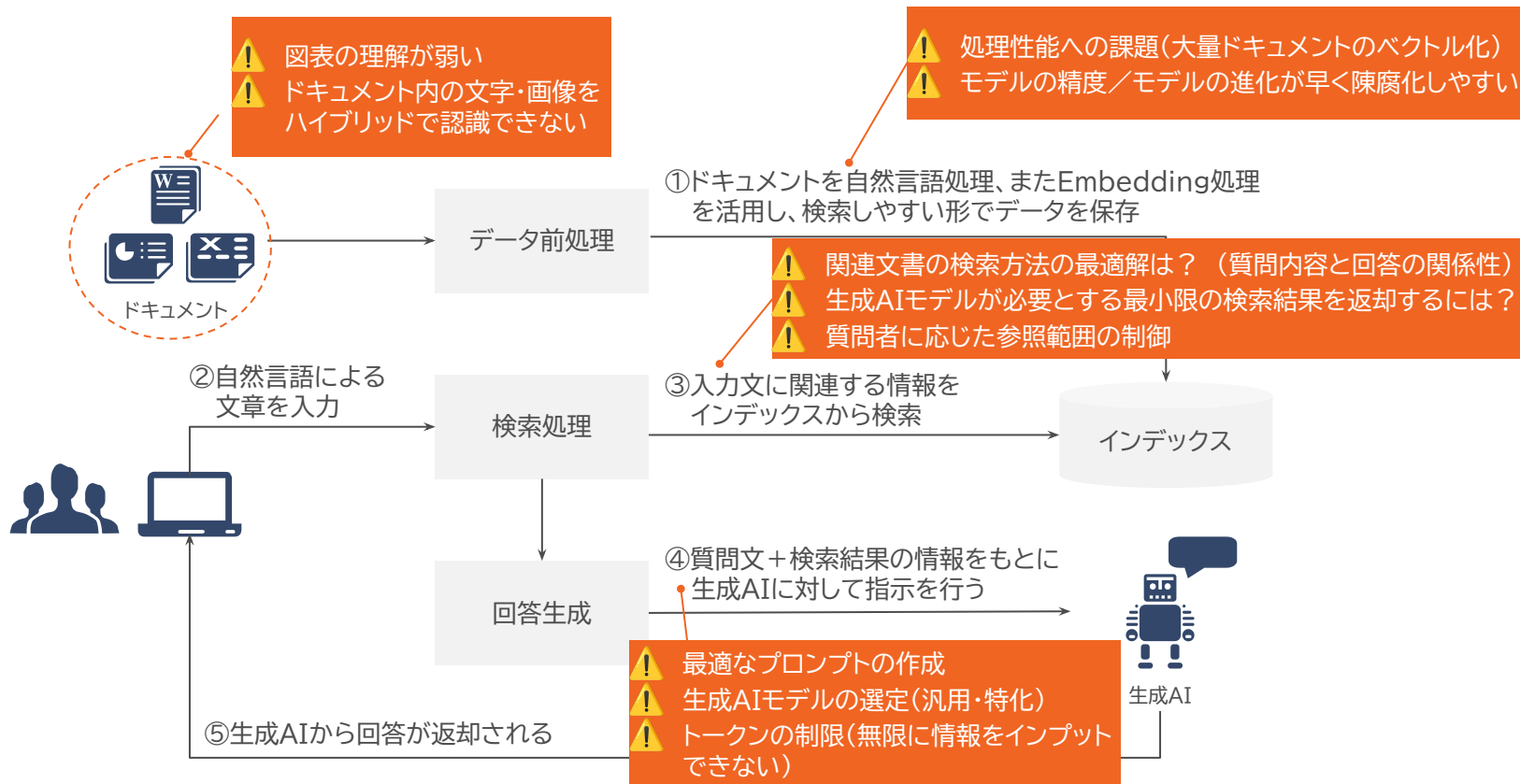
資料内の情報と質問それぞれの意味をスコア化し、近しい資料やデータから回答を生成するため、高い回答精度を実現

セキュ
リティ

オンプレ提供および
ファイル権限に応じた回答

Chat EIはクラウドサービスではなく、社内のオンプレミス環境に構築するためセキュアに運用可能。またファイルの閲覧権限に応じた回答生成により、情報の取り扱いも安心

RAGを構成する関連技術とよくある課題



RAGの回答が上手く返却できるケース

✓ 回答対象となる文書が文章で書かれている（※プレゼン資料ではない）



「電柱が倒壊するトラブルの対処法を教えてください」



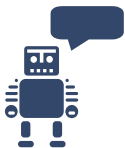
「旅費の精算方法を教えてください」

事故報告書

標記の件について、下記の通り報告いたします。

記

発生日時	2021年11月
発生場所	近畿
事故種別	電気工作物の破損等による物損
事故の状況	道路端の電柱（鉄筋コンクリート柱、1974年製）が地際で折損して家屋の敷地に倒壊し、家屋の外構フェンス、低圧引込線支持点の壁、自動販売機の引込小柱を損傷した。
被害の状況	道路端の電柱（鉄筋コンクリート柱、1974年製）が地際付近で折損・倒壊 ・家屋の外構フェンスの損傷 ・支持点付近の壁と引込小柱の損傷
事故発生種別	・荷重に影響がある設備変更の際に、確実に強度計算を行うよう再周知、再徹底する。 ・劣化様相（傾斜、湾曲、ひび割れ）が見られる電柱を調査し、不平衡荷重が強度を上回る箇所は、速やかに昇柱禁止・倒壊防止措置を講じ、改修を計画する。 ・不平衡荷重及び電柱強度の調査結果を踏まえ、健全性評価基準を見直しながら、継続的に適切な改修を進める。
事故の原因	原因は、経年的な不平衡荷重により、電柱下部に横ひび割れが発生し、ひび割れから水分が浸入して複数の鉄筋が腐食したことにより、鉄筋強度が低下したため倒壊に至ったと推定される。
今後の対策	・荷重に影響がある設備変更の際に、確実に強度計算を行うよう再周知、再徹底する。 ・劣化様相（傾斜、湾曲、ひび割れ）が見られる電柱を調査し、不平衡荷重が強度を上回る箇所は、速やかに昇柱禁止・倒壊防止措置を講じ、改修を計画する。 ・不平衡荷重及び電柱強度の調査結果を踏まえ、健全性評価基準を見直しながら、継続的に適切な改修を進める。



旅費規程

（目的）

第1条 この規程は、〇〇株式会社（以下「会社」という。）の役員および社員が就業規則第〇条および第〇条に基づく出張または転勤のため旅行するときの旅費に関する事項を定める。

2 この規程に定めのない事項については、そのつど定めるものとする。

（出張の定義）

第2条 この規程において出張とは、会社の出張命令により指示された往復〇km以上のものをいう。

2 出張に該当しない短距離移動に伴う交通費については、別に定める。

（旅費の種類）

第3条 旅費は、国内出張旅費と転勤旅費に区分する。

（旅費の計算）

第4条 旅費は、業務遂行上必要かつ最も経済的な経路（以下「順路」という。）によって計算する。ただし、業務の都合その他やむを得ない事由があると会社が承認したときは、この限りではない。

（承認）

第5条 出張または転勤のため旅行するとき、その日程および順路について、あらかじめ所属長の承認を得なければならない。

（旅費の前払い）

第6条 旅費は、原則として帰社後または赴任後精算し、支払うものとする。

2 前項の規定にかかわらず、会社の承認を得た場合は、概算による旅費の前払いを受けることができる。この場合、帰社後または赴任後〇日以内に精算しなければならない。

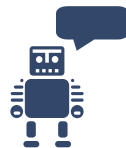
（旅費の精算）

第7条 旅費は、帰社後または赴任後〇日以内に所定の様式に必要な事項を記入して所属長の承認を受け、精算しなければならない。

2 長期出張の場合は、毎月〇日を精算日とする。

3 精算金は、毎月〇日までの精算分を当月給与支給日に支給する。

（旅費の不払い）



RAGで回答が上手く返却できないケース

! 表形式の内容



「電柱が倒壊するトラブルの対処法を教えてください」

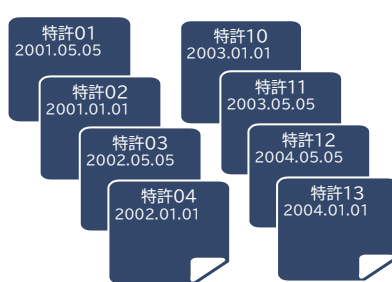
発生事象	発生原因	対処方法
当該高圧配電線路の配電柱上の電線接続箇所が過熱して、電線が断線・落下するとともに、スリーブカバーの焼損・溶融・滴下により、野火が発生して下草(笹等)を焼失したため、電気火災事故になった。(配電電圧:6600V)	当該高圧配電線の電線(銅線)とジャンパー線(アルミ線)を接続するアルミ線用圧縮スリーブ(分岐スリーブ)において、コンパウンドの充填不足によりスリーブ内に塩水が浸入してアルミ腐食生成物(アルミナ)が生成され、接触抵抗の増加・過熱により電線が断線・落下するとともに、スリーブカバーが焼損・溶融・滴下して下草(笹等)に引火したため、電気火災(野火)に至ったものと推定される。	・管内のアルミ線と銅線の接続箇所について、断線箇所と同一時期に施工した箇所(70箇所)の状態に異常が無いことを確認した。 ・非難燃性分岐スリーブカバーの取り付け箇所について、難燃性分岐スリーブカバーへの取替を実施した。 ・本事象について全道の事業所等に周知するとともに、電線接続時には、分岐スリーブ内にコンパウンドの必要量が充填されていることを確認し、コンパウンド充填が不十分な場合には、分岐スリーブを取替えて電線接続を行うように指導する。
道路端の電柱(鉄筋コンクリート柱、1974年製)が地盤沈下により、家屋の敷地に倒壊した。この外構フェンスは、高圧引込線支持点の自動販売機の引込小柱に設置した。	経年的な不平衡荷重により、横び割れが発生し、ひ水分が浸入して複数の鉄筋ごとにより、鉄筋強度が低下し、倒壊に至ったと推定される。	は、この工事に伴って、当該外構フェンスの改修を計画する。 ・不平衡荷重及び電柱強度の調査結果を踏まえ、健全性評価基準を見直しながら、継続的に適切な改修を進める。

表形式を理解しておらず、対処方法に記載されている内容のみから対処方法というコンテキストを判断するのが難しいケースがある

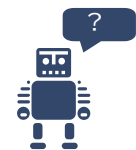
! 数多くの文書を跨いだ情報の集約



「2000年以降の特許件数を年ごとに集計して」



インプットとなる特許文書を検索で取得できたとしても、各文書に記載されている公開日の情報を取得しなければならず、多くの文書が対象になると生成AIへのインプットができない、もしくは、情報量が多すぎると、回答精度が悪くなる傾向にあり、難しい



! 専門用語(社内でのみ通用するワード)などが利用されているケース

! 質問内容に曖昧な内容があり、そもそも答えられない(人が空気を読んで回答している内容) など

適用分野
検討中

自動車部品メーカー

社内向けの教育素材への
問合せ効率化

検証内容

- ・ 社内研修向けの講義資料(約150ページ)の内容を解析
- ・ 想定QA10問に対して、想定回答例と実際の回答との類似度を5段階で判定

検証結果

- ・ 質問に対して精度良い回答が得られた(想定QAの回答例との平均類似度4.3を達成)
- ・ クローリング機能により定期的に取り込める点で運用性が高い点を評価

評価のみ

非鉄金属メーカー

社内の規定を対象とした
事務作業の効率化

検証内容

- ・ 社内の規定、および、契約締結のルール等の文書を解析
- ・ 想定QA8問に対する実際の回答を比較

検証結果

- ・ 契約書に関する問合せは、精度高く回答ができていた(想定QA8問に対して、6問が予想解答とほぼ一致)
- ・ 出典元の文書は正しく特定できているものの、金額による承認者の分岐で一部間違いが発生

評価のみ

自動車部品メーカー

保守報告書を用いた
製造ラインの保守作業効率化

検証内容

- ・ 200件のExcel形式の報告書をそのまま解析
- ・ 想定QA20問に対して、想定回答と実際の回答との類似度を5段階で判定

検証結果

- ・ 回答内容の精度はまちまち（平均類似度2.6）
- ・ 複数のファイルや一つのファイルの複数箇所を参照する必要がある場合は不得意な傾向

評価
利用中

自動車メーカー

過去のトラブル履歴を用いた
製造ラインの保守作業効率化

検証内容

- ・ 4,000件にも及ぶ、過去のトラブル履歴データを解析
- ・ 想定QA5問に対して、想定回答と実際の回答との類似度を5段階で判定

検証結果

- ・ 他のサービスでは満足のいく精度が得られなかったが、Chat EI では高い精度が得られたため、早期に現場へ展開し効果を検証中

EmbeddingとRAG

🔄めくる

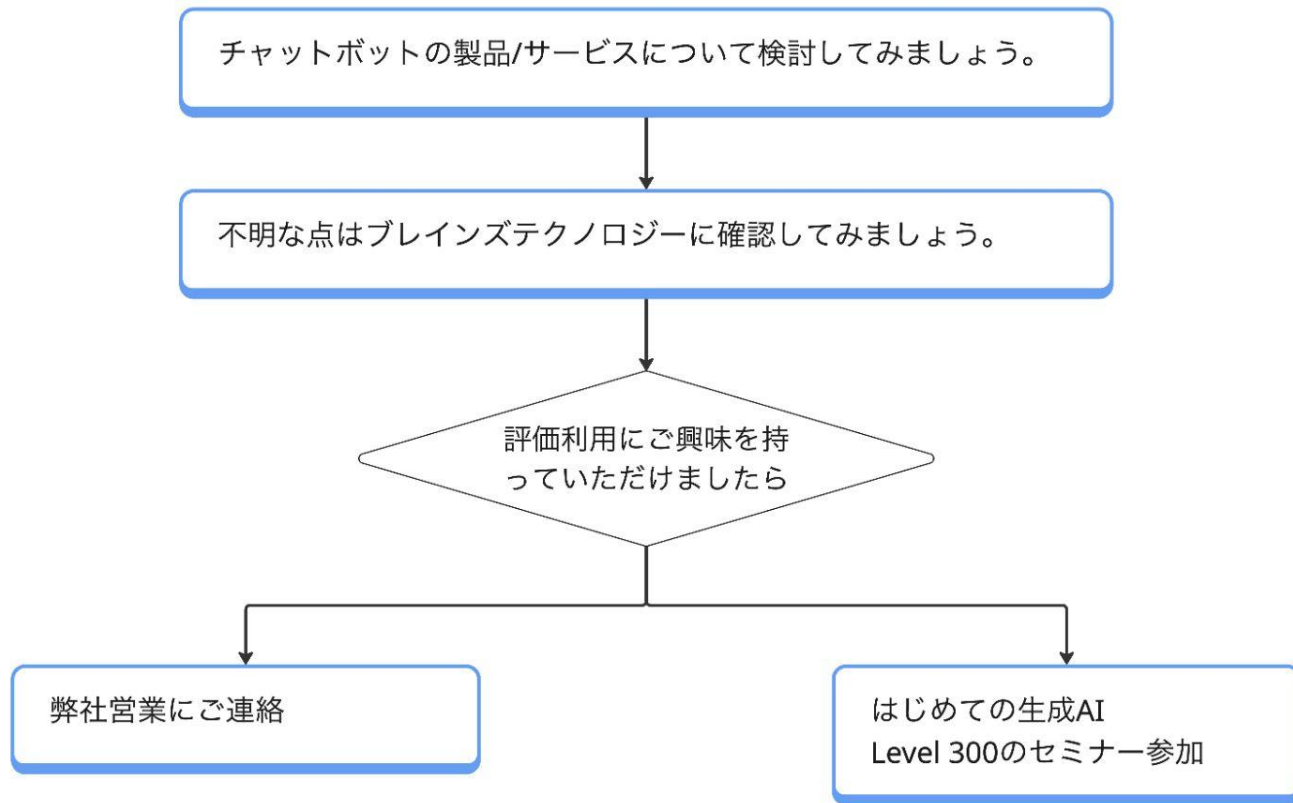
Chat EIの
システム構成

🔄めくる

よくある課題と
検証事例

🔄めくる

次のステップ



[www.brains-tec...](http://www.brains-tec.com)

Neuron

【セミナー】初めての生成AI Lv.300 | Neuron ES (公式)

開催日程
2025年7月3日 (木) 16:00~17:00
学べること
トライアルの流れ関係部署に説明できるよう【..】